

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(19)日本国特許庁 (JP)

(12) 特 許 公 報 (B 2)

(11)特許番号

第2834471号

(45)発行日 平成10年(1998)12月9日

(24)登録日 平成10年(1998)10月2日

(51)Int.Cl.⁶
G 1 0 L 3/00

識別記号
5 5 1
5 1 5
5 3 1

F I
G 1 0 L 3/00

5 5 1 E
5 1 5 B
5 3 1 K

請求項の数1(全4頁)

(21)出願番号 特願平1-97735

(22)出願日 平成1年(1989)4月17日

(65)公開番号 特開平2-275499

(43)公開日 平成2年(1990)11月9日
審査請求日 平成7年(1995)8月4日

特許法第30条第1項適用申請有り 電子情報通信学会技術研究報告【音声 SP88-155, P. 9~14 (平成元年3月)

(73)特許権者 99999999
日本電信電話株式会社
東京都新宿区西新宿3丁目19番2号
(72)発明者 三樹 駿
東京都千代田区内幸町1丁目1番6号
日本電信電話株式会社内
(72)発明者 浜田 洋
東京都千代田区内幸町1丁目1番6号
日本電信電話株式会社内
(72)発明者 中津 良平
東京都千代田区内幸町1丁目1番6号
日本電信電話株式会社内
(74)代理人 弁理士 草野 阜
審査官 榎本 剛

(56)参考文献 特許2704216 (JP, B 2)

最終頁に続く

(54)【発明の名称】 発音評価法

1

(57)【特許請求の範囲】

【請求項1】被評価話者の音声を標準話者の音声に話者適応化し、その話者適応化された被評価話者の音声と上記標準話者の音声とのフレームの対応をスペクトル類似性に基づいてとり、その後パラメータの平均値・分散値を用いて正規化した被評価話者と標準話者の韻律的特徴を上記対応づけたフレームごとに比較することを特徴とする発音評価法。

【発明の詳細な説明】

【産業上の利用分野】

この発明は評価対象の話者が外国語等を発声した時に、その発音の良さを評価する発音評価法に関するものである。

【従来の技術】

通常、外国語発音の韻律的な良さに関する定量的な評

2

価は、評価する外国語の発音が良い音声、例えばその外国語を母国語とする者（以後、標準話者と呼ぶ）の音声を基準とし、それと評価対象となる話者（以後、被評価話者と呼ぶ）における同じ発生内容の音声がどれほど近いか、という観点で行うのが一般的である。この場合、時間軸方向における発声の非線形な変動の違いを無視し、同じ音韻を発声している部分同士の韻律を比較できるようにするために、音声をスペクトルの変化がないと見なせる一定時間間隔（これをフレームと呼ぶ、以下同じ）に分け、フレーム単位でスペクトル類似性に基づく標準話者の音声と被評価話者の音声との時間対応をとった後、スペクトル特性が対応している、つまり音韻が一致していると考えられるフレーム同士で被評価話者と標準話者の韻律的特徴がどれほど異なっているかを比較し、評価するといった方法をとる。

[発明が解決しようとする課題]

この評価法において従来の技術では、上記フレーム対応を、標準話者と被評価話者の音声をそのままDPマッチング法、HMMなどでマッチングし、その対応結果によって定めていた。しかし、発声者が異なる音声に単純にDPマッチング法などを適応した場合、各音韻におけるスペクトル特性の個人差によって、被評価話者と標準話者の物理的スペクトル類似性によるフレーム対応が音韻対応とうまく一致しないことが多く、そのため音韻対応が誤った状態で批評価話者と標準話者との韻律的特徴の比較を行ってしまい、評価に悪影響を及ぼしていた。

また、韻律的特徴の比較において、従来の単純にそのパラメータ同士を比較する方法では、その平均基本周波数・平均音声パワーの個人差やダイナミックレンジの個人差等の個�性と、韻律パラメータの基本形状の差を分離して評価することができず、これも評価に悪影響を及ぼしていた。

この発明の目的は、従来の技術では不正確であったスペクトル類似性に基づく被評価話者と標準話者のフレーム対応による音韻対応の精度を上げ、より正確な韻律的特徴の評価を行えるようにし、かつ平均基本周波数・平均音声パワーの個人差やダイナミックレンジの個人差等の個�性と韻律パラメータの基本形状の差とを分離して評価することができる発音評価法を提供することにある。

[課題を解決するための手段]

この発明は、被評価話者の音声を例えればヒストグラムを用いたコードブックマッピングなどの手法により標準話者の音声に話者適応化し、その話者適応化した後の被評価話者の音声と標準話者の音声とのスペクトル類似性に基づいたフレーム対応を計算することによって、音韻対応する精度を上げ、その対応フレームごとに被評価話者と標準話者の韻律的特徴を比較することによって、発音の韻律的な良さのより正確な評価が行えることおよび被評価話者と標準話者の韻律的特徴をその平均値・分散値を用いて正規化した後、上記のフレーム対応ごとにその差異を比較することによって、韻律パラメータの基本形状の差のより正確な評価が行えることも最も主要な特徴とする。

[作用]

話者適応化によってスペクトル特性の個人差の関係がより明確になり、それによって個人差を吸収でき、スペクトル類似性に基づくフレーム対応手法による音韻対応精度が向上する。

また平均および分散を正規化することによって、平均基本周波数・平均音声パワーの個人差やダイナミックレンジの個人差等の個�性を吸収し、韻律パラメータの基本形状の差のみを分離評価できる。

[実施例]

第1図はこの発明の一実施例を説明する図であって、

被評価話者の発声した音声の韻律的な良さを評価するものである。

最初に標準話者と被評価話者が同一の音声セット（単語、単文等）を発声する。次にこの発声データを用いて被評価話者の音声を標準話者の音声に話者適応化する。ここで用いる方法を以下に示す。

まず、コードブック生成部1において話者ごとのコードブックを作成する。発声された標準話者の音声2を音声分析部3においてフレーム単位に分析する。分析手法

10 としては、バンドパスフィルタ分析、線形予測分析、FFT分析などが提案されているが、そのいずれかを用いて分析を行えばよい。ここでp次のLPCケプストラム係数をフレームのスペクトルに関する特徴パラメータとして用いる。次に分析後の音声をクラスタリング演算部4でクラスタリングし、標準話者音声の代表的なスペクトルパターンである、あらかじめ定められた数nのコードベクトルからなる標準話者のコードブック5を作成する。このクラスタリングの手法についてはLinde, Buzo, and Gray: "An Algorithm for Vector Quantizer Design", IEEE Trans. Comm., Vol. COM-28, 1980に詳しい。

被評価話者の音声6も同様な手続きで音声分析部7で分析し、クラスタリング演算部8でクラスタリングを行い、コードブック9を作成する。

次にコード列生成部10において上記で作成した標準話者のコードブック5を用いて、標準話者の音声2をベクトル量子化部11でフレーム単位にベクトル量子化し、ベクトルコード列12を作成する。被評価話者の音声6も同様にベクトル量子化部13でベクトル量子化し、ベクトルコード列14を作成する。

30 30 ここで話者適応化部15において、被評価話者の音声を標準話者の音声に適応化を行う。話者適応化については、いくつかの手法が提案されているがここではShikano, Lee, and Reddy: "Speaker Adaptation through Vector Quantization", Proc. ICASSP-86, 49.5, 1986で提案されたヒストグラムを用いた方法による例を示す。

まずマッチング演算部16において同一対象音声での標準話者のベクトルコード列12と被評価話者のベクトルコード列14との間でマッチング演算を行い、2つのベクトルコード列のフレーム対応を計算する。この対応を利用して、ヒストグラム生成部17において、被評価話者のコードブック内での個々のコードベクトルに対する標準話者のコードブック内でのコードベクトルの対応をヒストグラムの形で表す。そこで適応化コードブック生成部18において、話者適応化前の被評価話者の各コードベクトルに対応する話者適応化後のコードベクトルを、適応化前コードブックに対するヒストグラムを重みとして標準話者のコードベクトルを重み付き平均することによって作成する。そしてこの適応化後のコードベクトルを集め、話者適応化した被評価話者のコードブック19を作成する。

次にフレーム対応計算部20において、上記のようにして作成した適応化した被評価話者コードブック19内のコードベクトルと標準評価コードブック5内のコードベクトルを特徴パラメータとして用い、それぞれのベクトルコード列12, 14から上記コードブック5, 19を参照しながらフレーム間のスペクトル類似性を計算し、マッチング手法によりフレーム対応21を決定する。マッチング手法としてはDPマッチング法、HMMなど提案されているいずれかの方法が利用できる。DPマッチング法を用いたこのマッチング方法の報告が鹿野：“入力音声のベクトル量子化による単語音声認識”，音響学会音声研究会資料，S82-60, 1982になされている。

最後に韻律評価部22において、韻律の良さの評価値を計算する。

ここでは評価する韻律的特徴として基本周波数と音声パワーを用い、またその正規化は以下の方法をとることにする。まず、対象音声の基本周波数はピッチ抽出・スマージング部23において変形相関法などのピッチ抽出手法により抽出した後、倍・半ピッチなどの抽出エラーを除くためスマージングをかける。次にピッチ正規化部24において対数をとり、平均基本周波数の個人差を正規化するため抽出した基本周波数の有声部の平均を全体から引き、さらにダイナミックレンジの個人差を正規化するため同じく有声部の標準偏差で割る。音声パワーも同じくパワー計算部25により計算され、パワー正規化部26において対数をとった後、平均音声パワーを正規化するため音声パワーの音声区間の平均を全体から引き、さらにダイナミックレンジの個人差を正規化するため音声区間の標準偏差で割る。正規化はこのように、単純に平均・分散を一致させるだけでなく、人間の感覚にあわせてパラメータを非線形伸縮させる方法も可能である。

次に適応化後のフレーム対応21を基に韻律的特徴の比較を行う。ここでは、比較方法として差の絶対値を用いることとする。この他にも上記絶対値に重み付けを行

う、対応する韻律的特徴パラメータの相関をとる、など話者適応化後の被評価話者音声と標準話者音声のフレーム対応21に基づいた種々の比較方法が可能である。ここで正規化した基本周波数と音声パワーはそれぞれの韻律比較部すなわちピッチ比較部27、パワー比較部28において上記の方法で計算した対応フレームごとに差の絶対値を計算し、たとえば基本周波数は有声部のみで平均、音声パワーは音声区間全体で平均し評価値29, 30とする。この評価値29, 30は他の方法で計算した音韻的な評価と組み合わせて、総合的な発音の良さの評価値とすることが可能である。

このようにこの方法は、標準話者に話者適応化して個人性を吸収した後の被評価話者の音声と標準話者の音声とのフレーム対応を計算する構造になっているから、被評価話者と標準話者の音韻の対応がより正確にとれ、また、基本周波数の有声部での平均・分散、音声パワーの音声区間での平均・分散で正規化することにより、平均基本周波数・平均音声パワーの個人差やダイナミックレンジの個人差等の個人性が吸収できる。

20 その効果としては、より精度の高い音韻の対応に基づく韻律的特徴の比較が可能になり、かつ韻律パラメータの基本形状のみを分離評価できるため、より正確な評価ができる。

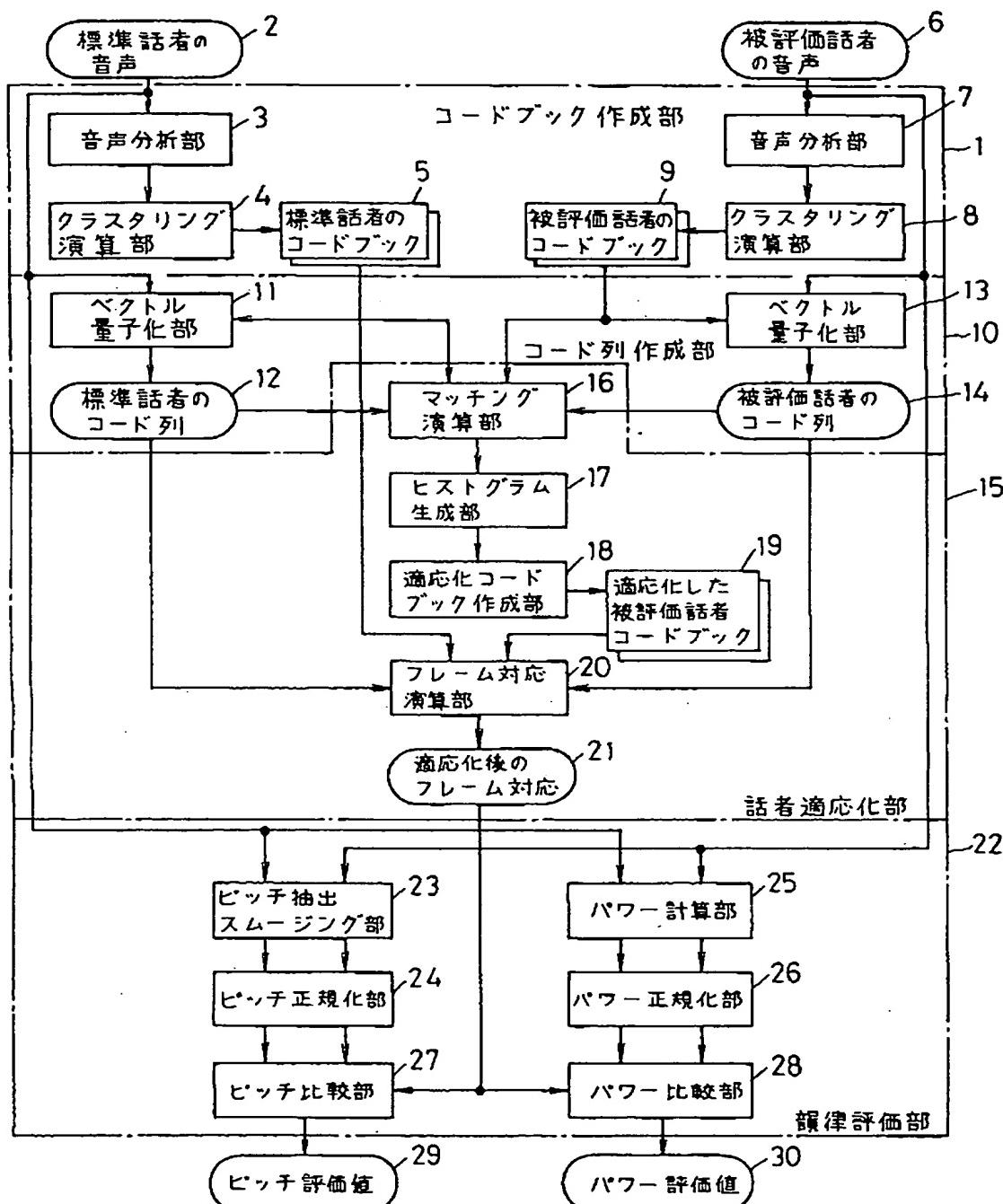
【発明の効果】

以上説明したように、この発明は話者適応化つまりスペクトル的な個人性を吸収した後の被評価話者の音声と標準話者の音声とのフレーム対応をとることにより、音韻の対応精度が向上し、韻律的特徴評価がより正確になり、かつ、韻律パラメータをその平均値・分散値を用いて正規化することにより、韻律的特徴の基本形状のみを分離評価できる利点がある。

【図面の簡単な説明】

第1図はこの発明による発音評価法の実施例を示すプロック図である。

【第1図】



フロントページの続き

(58)調査した分野(Int.Cl.®, DB名)

G10L 3/00 551

J I C S T ファイル (J O I S)